

# ON INTEGERS WHICH ARE REPRESENTABLE AS SUMS OF LARGE SQUARES

ALESSIO MOSCARIELLO

**ABSTRACT.** We prove that the greatest positive integer that is not expressible as a linear combination with integer coefficients of elements of the set  $\{n^2, (n+1)^2, \dots\}$  is asymptotically  $O(n^2)$ , verifying thus a conjecture of Dutch and Rickett. Furthermore we ask a question on the representation of integers as sum of four large squares.

## INTRODUCTION

Additive number theory is the branch of arithmetic that studies particular subsets of integers (such as arithmetic progressions, prime numbers or squares) and their behavior under addition. A typical class of problems in this area concerns determining which positive integers can be expressed as sums of elements of a fixed subset of integers. In this aspect, one of the earliest results of additive number theory is Lagrange's Four Square Theorem, stating that every non-negative integer can be expressed as sum of four squares. Another classical problem is the Diophantine Frobenius Problem (cf. [6]), that asks for the greatest integer that is not representable as a linear combination with non-negative integer coefficients of elements of a subset of integers. In connection to both topics, it is interesting to understand which numbers are representable as sum of large squares. In this work, we investigate the largest number that cannot be represented as a sum of squares of integers greater than a fixed  $n$ . In particular, we answer affirmatively a conjecture of Dutch and Rickett (cf. [1]), stating that the greatest integer that is not expressible as a sum of squares of integers greater than or equal to  $n$  grows like  $n^2$ , namely we prove the following result:

**Theorem 1.** *Let  $\Gamma_n$  be the set of integers that can be expressed as a finite sum of squares of integers greater than or equal to  $n$ , and denote by  $F(\Gamma_n)$  the greatest integer not belonging to  $\Gamma_n$ . Then  $F(\Gamma_n) = O(n^2)$ .*

Finally we ask a question related to the representation of an integer as a sum of four large squares.

## 1. MAIN RESULT

Denote by  $\mathbb{N}$  the set of non-negative integers. Since  $n^2$  and  $(n+1)^2$  are relatively prime for any  $n \in \mathbb{N}$ ,  $n \neq 0$ , then Sylvester's formula (cf. [7]) states that every integer greater than or equal to  $(n^2 - 1)[(n^2 + 1) - 1]$  can be written as a non-negative linear combination of  $n^2$  and  $(n+1)^2$ , thus the set  $\mathbb{N} \setminus \Gamma_n$  is finite. Recently it has been proved (cf. [5]) that the greatest integer that is not representable as a non-negative linear combination of  $n^2, (n+1)^2, (n+2)^2$

---

2010 *Mathematics Subject Classification.* 11D07, 11P05.

*Key words and phrases.* Frobenius number, Four Square Theorem.

is of order  $O(n^3)$ . Concerning our problem, it is known that  $F(\Gamma_n) = o(n^{2+\epsilon})$  for any  $\epsilon > 0$  (cf. [1]). We estimate  $F(\Gamma_n)$  by focusing on representations that involve the least possible number of squares. The well-known Four Square Theorem, due to Lagrange, states that every positive integer can be written as a sum of four squares. Furthermore, if we denote by

$$r(n) := \#\{(a_1, a_2, a_3, a_4) \in \mathbb{Z}^4 \mid a_1^2 + a_2^2 + a_3^2 + a_4^2 = n\}$$

and by  $\sigma'(n)$  the sum of divisors of  $n$  that are not divisible by 4, in 1828 Jacobi proved that  $r(n)$  is equal to  $8\sigma'(n)$  (cf. [4]). Let  $\sigma(n)$  stand for the sum of divisors of  $n$ . Notice that if  $n \in \mathbb{N}$  is such that  $n \not\equiv 0 \pmod{4}$  then  $\sigma'(n) = \sigma(n) > n$ , hence for  $n \rightarrow +\infty$ ,  $n \not\equiv 0 \pmod{4}$  we get that  $r(n) \rightarrow +\infty$ . Therefore it makes sense to ask for representations as sum of four relatively large squares. However, the elements of the infinite set  $\mathcal{B} = \{1, 3, 5, 9, 11, 17, 29, 41\} \cup \{4^\alpha 2, 4^\alpha 6, 4^\alpha 14 \mid \alpha \in \mathbb{N}\}$  cannot be expressed as sum of four non-zero squares (see [2], Chapter 6, Theorem 3). Taking that into account, we pose the following question:

**Question 2.** *Is there  $K$  in  $\mathbb{N}$  such that for every  $n$  in  $\mathbb{N}$  there exists  $a_1, \dots, a_4$  such that  $n = a_1^2 + a_2^2 + a_3^2 + a_4^2$  and for each  $i$  either  $a_i = 0$  or  $a_i \geq \frac{\sqrt{n}}{K}$ ?*

Note that if such  $K$  exists then, for every  $n \in \mathbb{N}$ , every positive integer  $M \geq K^2 n^2$  is a sum of four squares contained in the set  $\{0, n^2, (n+1)^2, \dots\}$ : hence  $M \in \Gamma_n$ ,  $F(\Gamma_n) \leq K^2 n^2$  and  $F(\Gamma_n) = O(n^2)$ . It is easy to see that such a  $K$  cannot be less than 8 (by taking  $N = 55$ ), and computational evidence suggests that  $K = 8$  affirmatively answers Question 2. We are going to give a partial answer to this question by proving that such a  $K$  must exist, in order to prove our main theorem. Consider now the 4-dimensional sphere of radius  $\sqrt{n}$

$$\mathcal{S}_n := \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_3^2 + x_4^2 = n\}.$$

The representations of  $n$  as sum of four integer squares correspond to the integral points of  $\mathcal{S}_n$ . Question 2 is thus related to the distribution of these integral points over the sphere  $\mathcal{S}_n$ . This problem has been widely studied in literature in the more general context of quadratic forms: it has been proved (cf. [3], Theorem 11.5) that for an arbitrary quadratic form  $Q$  if  $r \geq 4$  then the vectors  $a \in \mathbb{Z}^r$  with  $Q(a) = n$  are asymptotically equidistributed as  $n \rightarrow \infty$  over integers satisfying the congruence  $Q(a) \equiv n \pmod{128(\det(Q))^3}$ . We will now explain the meaning of the words “asymptotically equidistributed”. First, points on  $\mathcal{S}_n$  scale concentrically to points of the unit sphere  $\mathcal{S}_1$ . Thus we can study the distribution of integral points over  $\mathcal{S}_n$  by considering the distribution of the associated scaled points of  $\mathcal{S}_1$ . For this purpose we define, for a domain  $F \subseteq \mathcal{S}_1$ ,

$$r_F(n) := \#\left\{(a_1, a_2, a_3, a_4) \in \mathbb{Z}^4 \mid (a_1, a_2, a_3, a_4) \in \mathcal{S}_n, \left(\frac{a_1}{\sqrt{n}}, \frac{a_2}{\sqrt{n}}, \frac{a_3}{\sqrt{n}}, \frac{a_4}{\sqrt{n}}\right) \in F\right\},$$

that is the number of integral points of  $\mathcal{S}_n$  that scale to points of  $F$ . With this notation, by “asymptotically equidistributed” we mean that if  $\mathcal{F}$  is a convex domain with smooth boundaries on  $\mathcal{S}_1$ , then  $r_{\mathcal{F}}(n) \sim |\mathcal{F}|r(n)$ , where  $|\mathcal{F}|$  denotes the normalized surface area of  $\mathcal{F}$  (assuming  $|\mathcal{S}_1| = 1$ ). Notice that, if we consider the quadratic form  $Q(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2$ , then by Lagrange’s Four Square Theorem for every integer  $n$  there exists  $a = (a_1, a_2, a_3, a_4) \in \mathbb{Z}^4$  such that  $Q(a) = n$ , thus the congruence  $Q(x) \equiv n \pmod{128}$  admits solutions for every  $n \in \mathbb{N}$ ; this consideration allows us to discard this constraint in the following.

This lemma partially answers Question 2:

**Lemma 3.** *Let  $n$  be a squarefree integer. Then, for  $n$  sufficiently large, there is an integral point  $(a_1, a_2, a_3, a_4) \in \mathcal{S}_n$  such that  $a_i \geq \frac{\sqrt{n}}{8}$ .*

*Proof.* Notice first that for  $n$  squarefree we have  $n \not\equiv 0 \pmod{4}$ , hence  $\sigma(n) = \sigma'(n)$  and  $r(n) = 8\sigma(n) \geq 8(n+1)$ . Considering the domain

$$\mathcal{F} := \left\{ (x_1, x_2, x_3, x_4) \in \mathcal{S}_n \mid x_i \geq \frac{1}{8} \text{ for every } i = 1, 2, 3, 4 \right\} \subseteq \mathcal{S}_1,$$

we have  $|\mathcal{F}| > 0$ . Thus  $r_{\mathcal{F}}(n) \sim |\mathcal{F}|r(n) \geq |\mathcal{F}|8(n+1)$ , therefore on squarefree integers the sequence  $r_{\mathcal{F}}(n)$  is asymptotic to a sequence that is larger than  $|\mathcal{F}|8n$ , so the values of  $r_{\mathcal{F}}(n)$  must be strictly positive for sufficiently large squarefree  $n$ , and this implies our thesis.  $\square$

A consequence of Lemma 3 is:

**Lemma 4.** *There exists a positive constant  $K$  such that for every squarefree positive integer  $n$  there is an integral point  $(a_1, a_2, a_3, a_4) \in \mathcal{S}_n$  such that if  $a_i \neq 0$  then  $a_i \geq \frac{\sqrt{n}}{K}$ .*

*Proof.* By Lemma 3 there exists  $N \in \mathbb{N}$  such that for every squarefree positive integer  $n \geq N$  there is an integral point  $(a_1, a_2, a_3, a_4) \in \mathcal{S}_n$  with  $a_i \geq \frac{\sqrt{n}}{8}$ . Therefore there is at most a finite set  $\{n_1, \dots, n_t\}$  of squarefree positive integers for which that property should not hold. However, by Lagrange's Four Square Theorem there should be an integral point  $(a_{i1}, a_{i2}, a_{i3}, a_{i4})$  on the sphere  $\mathcal{S}_{n_i}$  for every  $i = 1, \dots, t$ . Define now for every  $i = 1, \dots, t$  the constant

$$K_i := \max \left\{ \frac{\sqrt{n_i}}{a_{ij}} \mid a_{ij} \neq 0, j = 1, \dots, 4 \right\}.$$

Let now  $K = \max\{K_1, \dots, K_t, 8\} + 1$ . Then for every squarefree positive integer  $n$  we have that there is an integral point  $(a_1, a_2, a_3, a_4) \in \mathcal{S}_n$  such that whenever  $a_i \neq 0$  we have  $a_i \geq \frac{\sqrt{n}}{\max\{K_1, \dots, K_t, 8\}} > \frac{\sqrt{n}}{K}$ , thus concluding our proof.  $\square$

Now we are ready to prove the main theorem.

*Proof of Theorem 1.* Let  $n \in \mathbb{N}$ , let  $K$  be as in Lemma 4 and take  $n' \geq K^2 n^2$ . Suppose that  $n' = \rho^2 q$  where  $\rho, q \in \mathbb{Z}^+$  and  $q$  is squarefree. By Lemma 4 there is an integral point  $(a_1, a_2, a_3, a_4)$  on the sphere  $\mathcal{S}_q$  such that if  $a_i \neq 0$  we must have  $a_i \geq \frac{\sqrt{q}}{K}$ . Since  $a_1^2 + a_2^2 + a_3^2 + a_4^2 = q$  we have  $(\rho a_1)^2 + (\rho a_2)^2 + (\rho a_3)^2 + (\rho a_4)^2 = n'$ , and  $\rho a_i \geq \frac{\rho \sqrt{q}}{K} = \frac{\sqrt{n'}}{K} \geq n$  whenever  $\rho a_i \neq 0$ . Thus  $n' \in \Gamma_n$ . This implies  $F(\Gamma_n) < K^2 n^2$ , hence  $F(\Gamma_n) = O(n^2)$ .  $\square$

## 2. ASYMPTOTICS AND COMPUTATIONS

Theorem 1 states the existence of a bound for  $F(\Gamma_n)$ : however, in the proof it is not explicit how large the *sufficiently large*  $n$  should be to satisfy Lemma 3, and thus we cannot define explicitly the constant  $K$  of Lemma 4, although the existence of such  $K$  is sufficient for proving our main result. Here, we provide computational evidence that Question 2 might be satisfied by considering  $K = 8$ .

**Proposition 5.** *If there is an even positive integer  $n$  such that  $\tilde{K}$  is the minimal value of  $K$  such that Question 2 holds, then there is an infinite set of positive integers such that  $\tilde{K}$  is the minimal value of  $K$  such that Question 2 holds.*

*Proof.* Consider the integer  $4n$ .

Clearly, if  $a_1^2 + a_2^2 + a_3^2 + a_4^2 = n$  then  $(2a_1)^2 + (2a_2)^2 + (2a_3)^2 + (2a_4)^2 = 4n$  and  $2a_i \geq \frac{\sqrt{4n}}{K}$ , thus Question 2 holds for  $\tilde{K}$ . Suppose now that there exist  $K' \leq \tilde{K}$  and  $b = (b_1, b_2, b_3, b_4) \in \mathcal{Z}_{4n}$  such that  $b_i \geq \frac{\sqrt{4n}}{K'} = \frac{2\sqrt{n}}{K'}$  whenever  $b_i \neq 0$ . Since  $b_1^2 + b_2^2 + b_3^2 + b_4^2 = 4n$  and  $n$  is divisible by 8, then every  $b_i$  is even, and thus, considering  $a_i = \frac{b_i}{2}$  for every  $i = 1, 2, 3, 4$ , we have  $(a_1, a_2, a_3, a_4) \in \mathcal{Z}_n$  and  $a_i \geq \frac{\sqrt{n}}{K'}$  whenever  $a_i \neq 0$ . However,  $K' \leq \tilde{K}$ , thus leading to a contradiction.

Therefore,  $\tilde{K}$  is the minimum value of  $K$  such that Question 2 holds for  $4n$ . Since  $4n$  is even, this implies that  $\tilde{K}$  is the minimum value of  $K$  such that Question 2 holds for every integer of the form  $4^\alpha n$ ,  $\alpha \geq 0$ , and thus our thesis is proved.  $\square$

In light of this, we found the minimal integer constant  $K$  that work for small values of  $n$ . For this purpose, considering the set  $\mathcal{Z}_n$  of integral points of  $\mathcal{S}_n$ , we define, for each point  $a = (a_1, a_2, a_3, a_4) \in \mathcal{Z}_n$ ,  $L_a = \min\{a_i \mid a_i \neq 0\}$ . We can define a partial order  $\leq_L$  on  $\mathcal{Z}_n$  such that  $a \leq_L b$  if and only if  $L_a \leq L_b$  for every  $a, b \in \mathcal{Z}_n$ . In order to get the minimal value of  $K$  for a specific  $n$ , we need to find the maximal points of  $(\mathcal{Z}_n, \leq_L)$ . We computed with GAP these maximal points for  $n \leq 2560000$ . In Table 1,  $I_K$  denotes the set of integers  $n \leq 2560000$  such that  $K$  is the minimum value such that Question 2 holds, and  $S_K$  denotes the set of squarefree integers  $n \leq 2560000$  such that  $K$  is the minimum value such that Question 2 holds.

$K$	$\#I_K$	$\#S_K$	$\max S_K$
1	1600	1	1
2	1996562	1248201	2559998
3	561563	308003	2559999
4	232	59	1327
5	21	7	151
6	11	3	239
7	9	2	46
8	2	2	55

TABLE 1

From these computation it emerges that Question 2 holds with  $K = 8$  for every  $n \leq 2560000$ , and further considerations on this data suggest that  $K = 8$  might be a good choice for every integer  $n$ . Moreover, it seems that for the vast majority of integers  $n$  computed, Question 2 holds with  $K \leq 3$ . However, since  $46 \in I_7$ ,  $30 \in I_6$ ,  $78 \in I_5$  and  $10 \in I_4$ , Proposition 5 assures that  $I_4, I_5, I_6, I_7$  are all infinite sets; thus, surprisingly, even if we want to prove Question 2 for large values of  $n$  the constant  $K$  required cannot be lower than 7.

From Table 1 it seems that for large squarefree integers the minimum constant  $K$  required is either 2 or 3. Obviously  $K$  cannot be lower than 3, since if we consider squarefree integers  $n$  such that  $n \equiv 7 \pmod{8}$ , thus for every  $a = (a_1, a_2, a_3, a_4)$  such that  $a_1^2 + a_2^2 + a_3^2 + a_4^2 = n$  we have  $a_i \neq 0$  for every  $i = 1, 2, 3, 4$  and thus the minimum  $a_i$  cannot be greater than  $\frac{\sqrt{n}}{2}$ .

Nevertheless, it might be interesting to see whether Question 2 holds for squarefree integers with  $K = 3$ .

Solving this Question would provide bounds for  $F(\Gamma_n)$ . Denoting with  $F(n)$  the greatest positive integer that could not be represented as a sum of at most four squares of integers not lower than  $n$ , proving that Question 2 holds for every integer with  $K = 8$  would imply that  $F(\Gamma_n) \leq F(n) \leq 64n^2$ . Because of the restriction on the number of squares involved in our sum, we cannot verify directly these bounds for  $F(n)$ . However, under the assumption that Question 2 holds with  $K = 8$  for every positive integer, we can derive from the maximal vectors of positive integers up to  $2560000 = 200^2 \cdot 64$  a table of the various  $F(n)$  with  $2 \leq n \leq 200$ . Table 2 compares these values of  $F(n)$  with the relative values of  $F(\Gamma_n)$ . As Table 2 shows, for  $5 \leq n \leq 200$  it resulted that  $F(n) = 46 \cdot 4^{\lceil \log_2 n \rceil - 1}$ . Therefore, for these values it is easy to see that  $F(n) \leq 46n^2$ , as Table 2 and Figure 1 show.

$n$	$F(\Gamma_n)$	$F(n)$
2	23	55
3	87	184
4	119	239
5	201	736
6	312	736
7	376	736
8	455	736
9	616	2944
10	760	2944
20	2764	11776
30	5523	11776
40	9856	47104
50	15232	47104
60	21408	47104
70	28192	188416
80	36448	188416
90	46048	188416
100	57408	188416
125	84992	188416
150	122880	753664
175	164864	753664
200	215040	753664

TABLE 2

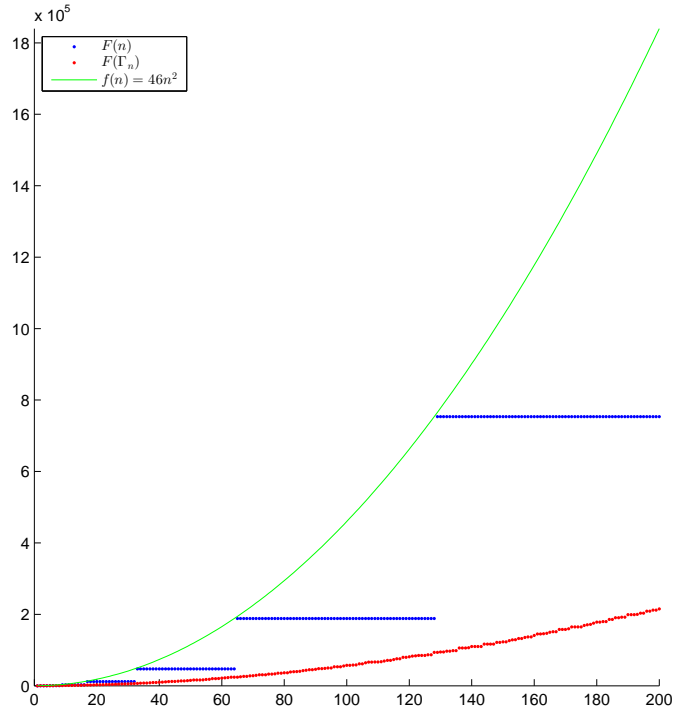


FIGURE 1

#### ACKNOWLEDGEMENTS

I would like to thank Alessio Sammartano for introducing me to this problem, and for his helpful comments and suggestions on this work. I would also like to thank the referee for his useful remarks on Lemma 3 and Section 2.

## REFERENCES

1. Dutch, K., Rickett, C.: *Conductors for sets of large integer squares*, Notes on Number Theory and Discrete Mathematics, Vol. 18 (2012), No. 1, 1621.
2. Grosswald, E.: *Representations of integers as sums of squares*, Springer-Verlag, (1984).
3. Iwaniec, H.: *Topics in Classic Automorphic Forms*, Graduate Studies in Mathematics, Vol. 17 (1997).
4. Jacobi, C.G.J.: *Note sur la dcomposition d'un nombre donn en quatre carrs*, J. Reine Angew. Math 3 (1828),191.
5. M. Lepilov, J. O'Rourke, I. Swanson, *Frobenius numbers of numerical semigroups generated by three consecutive squares or cubes*, Semigroup Forum, to appear (2015).
6. J. L. Ramírez Alfonsín, *The Diophantine Frobenius Problem*, Oxford University Press, Oxford (2005).
7. J. J. Sylvester, *Mathematical questions with their solutions*, Educational Times, 41 (1884).

(Alessio Moscariello) DIPARTIMENTO DI MATEMATICA E INFORMATICA, UNIVERSITÀ DI CATANIA, VIALE ANDREA DORIA 6, 95125 CATANIA, ITALY.

(Alessio Moscariello) SCUOLA SUPERIORE DI CATANIA, UNIVERSITÀ DI CATANIA, VIA VALDISAVOIA 9, 95125 CATANIA, ITALY.

*E-mail address:* `alessio.moscariello@studium.unict.it`